quarter, respectively, of the arcelin concentration present in SARC1-7.

At the lowest level tested, arcelin had no significant antibiosis effect on larvae. At the intermediate arcelin level, there was a significant increase in the larval life cycle duration but no significant effect on percentage emergence. However, the response of insects to the highest level of arcelin was nearly identical to that of insects reared on artificial seeds of SARC1-7 for both measures of resistance. This indicates that the presence of arcelin-1 in bean seeds confers resistance to *Z. subfasciatus*. The dosage response of larvae as measured by life cycle duration was nearly linear over the range of arcelin levels tested. For percentage emergence, a significant dosage response was observed only at the highest arcelin level, indicating that high levels are needed to affect this parameter.

TABLE V

Levels of resistance to *Z. subfasciatus* in intact and 'artificial' bean seeds with and without the addition of purified arcelin-1

Material screened	No. of days until adult emergence	Percentage emergence
Intact seed		
L12-56	32.2 (±0.6)	100.0 (±0)
Sanilac	31.3 (±0.6)	95.9 (±4.9)
SARC1-7	50.3 (±3.5)	7.4 (±7.4)
Calima	31.5 (±0.5)	93.0 (±5.7)
Artificial seed		
L12-56	38.4 (±1.5)	74.7 (±18.7)
Sanilac	37.8 (±2.0)	86.1 (±5.9)
SARC1-7	53.8 (±2.3)	18.4 (±14.7)
Calima	37.9 (±0.6)	87.7 (±13.6)
Sanilac + 2.5% arcelin-1	38.9 (±1.6)	76.1 (+11.8)
Sanilac + 5.0% arcelin-1	44.7 (±1.7)	76.1 (±19.5)
Sanilac + 10.0% arcelin-1	53.4 (±3.5)	18.4 (±17.6)

III. Cloning of Arcelin DNA Sequence

The following example provides a protocol for cloning arcelin-encoding DNA sequences, such as cDNAs.

Developing bean seeds of SARC1-7 were harvested 13–19 days after flouring and mRNA was isolated using the procedure described in Hall et al. (1978) *Proc. Natl. Acad. Sci. USA* 75:3196, except that the sucrose gradient centrifugations were omitted. The mRNA was used to construct a cDNA library in the pARC7 cDNA cloning vector. Alexander et al. (1984) *Gene* 31:79.

Candidate clones for arcelin-1 were selected by differential hybridization of colony lifts. Filters were prepared according to Taub et al. (1982) *Anal. Biochem.* 126:222. Filters were probed sequentially with three different probes. The probes, which were ^{32}P-labeled cDNA first strands made from mRNA fractions of developing seeds, were from (i) L12-56, a lectin-deficient bean line with phaseolin as the only major seed protein, (ii) the cultivar Sanilac, which contains PHA as well as phaseolin, and (iii) SARC1-7, which contains phaseolin, PHA, and arcelin-1. Arcelin cDNA candidates were selected as those colonies which were heavily labeled with SARC1-7 cDNA, but not by the other two cDNAs.

The nucleotide sequence (Sanger et al., 1978, FEBS Lett. 87:107) of candidate clone pAR1-11 contained multiple potential initiation sites near the 5' end of the clone, two of which were in the same large open reading frame encoding 269 amino acids (FIG. 1). Initiation at the third ATG would yield a 265 amino acid polypeptide. Comparison of this derived amino acid sequence to the N-terminal amino acid sequence of purified arcelin-1 protein demonstrated that pAR1-11 encodes arcelin-1. The mature protein sequence, which begins at codon 22 of the 265 amino acid open reading frame, matches the predicted sequence exactly through 47 of 48 amino acids determined by Edman degradation sequencing; amino acid residue number 12 yielded a blank by protein sequencing, presumably because of glycosylation of this predicted aspargine residue. The 21 residue peptide not found in the mature protein has the properties of a classic 'signal' peptide.

A clone of pAR1-11 in *E. coli* has been deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Md. 20852 USA. This deposit will be maintained under the terms of the Budapest Treaty.

Variations on the above embodiments are within the ability of one of ordinary skill in the art, and such variations do not depart from the scope of the present invention as described in the following claims.

We claim:

1. A composition comprising DNA molecules containing a coding sequence encoding arcelin-1 or an allelic variant thereof, said composition being substantially free of DNA molecules that do not contain said coding sequence.

2. A composition according to claim 1 wherein said DNA molecules are replicons.

3. A DNA molecule containing a coding sequence encoding arcelin-1 or an allelic variant thereof wherein said coding sequence is flanked by heterologous sequences which are T-DNA.

4. A DNA molecule containing a coding sequence encoding arcelin-1 or an allelic variant thereof wherein said coding sequence is flanked by heterologous sequences which are plant virus nucleic acid sequences.

5. A plant cell comprising a DNA sequence according to claim 3.

6. A plant cell comprising a DNA sequence according to claim 4.

7. A recombinant DNA sequence containing an expression cassette, said cassette containing a coding sequence encoding arcelin-1 or an allelic variant thereof, said coding sequence being under the transcriptional and/or translational control of regulatory sequences which are heterologous to said arcelin-encoding DNA.

8. A composition of cloning vectors, wherein each cloning vector in said composition contains a segment of DNA encoding arcelin-1 or the allelic variants thereof.

9. The recombinant DNA sequence of claim 7 wherein the control sequences of the expression cassette include a promoter selected from the group consisting of the nopaline synthase promoter, the octopine synthase promoter, the pea small subuit RUBP carboxylase promoter, and the soybean small subunit RUBP carboxylase promoter, the maize zein promoter, the wheat chlorophyll A/B binding protein promoter, soybean 7S-alpha'-conglycinin promoter, the soybean glycinin G2 promoter, soybean heat shock promoter, and the french bean phaseolin promoter.